Available online at www.sciencedirect.com

ScienceDirect

European Journal of Combinatorics 27 (2006) 1186–1198

European Journal of Combinatorics

www.elsevier.com/locate/ejc

# On threshold properties of $k$-SAT: An additive viewpoint

Alain Plagne

*Centre de Mathématiques Laurent Schwartz, UMR 7640 du CNRS, École Polytechnique, 91128 Palaiseau, France*

Available online 14 July 2006

**Abstract**

The SAT problem is one of the basic problems from complexity theory. When SAT is restricted to clauses of length $k$, we obtain the so-called $k$-SAT problem. For $k \geq 3$, it is a $\mathcal{NP}$-complete problem but it is believed that most of the instances are easy to solve. In fact, numerical evidence shows that a threshold phenomenon is to be expected. Up to now only upper bounds and lower bounds of the prospective value of the transition point have been obtained. Concerning lower bounds, they are obtained by considering special algorithms for which we can prove that it solves almost all instances of $k$-SAT if the ratio of the number of clauses by the number of variables is less than some given value.

In this expository paper, we propose a completely new approach on the problem of evaluating from below the prospective value of the transition point by showing a connection between $k$-SAT and number theory. More precisely, it is based on additive number theoretic considerations and avoids the use of any specific algorithm by directly counting the number of solutions to a system encoding an instance of $k$-SAT.


## 1. Introduction

The SAT problem is considered as a central one in complexity theory, see [17,20]. In propositional calculus, a SAT formula is composed of a set of $m$ clauses, each clause being a disjunction of literals (a variable or its negation) over a set of $n$ boolean variables (this is the Conjunctive Normal Form). The SAT problem is defined as, being given a SAT formula, determining whether there exists or not an assignment of the $n$ variables such that the $m$ clauses are satisfied simultaneously. If such a choice exists, the formula is said to be satisfiable.

*E-mail address:* plagne@math.polytechnique.fr.

0195-6698/$ - see front matter 
doi:10.1016/j.ejc.2006.06.014

The importance of the SAT problem is due to the fact that it is known to be $\mathcal{NP}$-complete and, even, the canonical example of $\mathcal{NP}$-completeness. It is a key for understanding computational issues. Practically, it has many advantages, among which that of being simple (the most basic intractable problem) and being fitted to computations and practical experimentations.

Even though being $\mathcal{NP}$-complete, there are many SAT formulae which are easily solvable: in other words, many SAT formulae are easy to determine satisfiable or not (the importance of this remark is clear if we think to the $\mathcal{P} = \mathcal{NP}$ problem). Restrictions of the SAT problem are thus of special interest.

As usual, we define a $k$-SAT formula to be a SAT formula such that the number of variables truly involved in each of the $m$ clauses is exactly $k$ (taken from the $n$ possible variables): in other words, each clause has length $k$. With this definition, the phenomenon that we just underlined appears clearly in the following manner: let us take $k = 2$ and define

$$\tau = m/n$$

then it is known that if $\tau < 1$ then almost every 2-SAT formula is satisfiable and if $\tau > 1$ then almost every 2-SAT formula is unsatisfiable. Here the "almost all" is to be understood with its probabilistic meaning that is, for given $n$ and $m = [\tau n]$, consider all the formulae in $n$ variables with $m$ clauses having the 2-SAT property, then "almost all" means with probability 1 as $n$ tends to infinity.

This is known as a threshold phenomenon. It is especially important to know when such a phenomenon does happen because it means that a unique type of instances concentrates the whole difficulty of the problem. Roughly speaking, the problem is either trivially solvable or without solution, except in one point (the transition point). This means that the complexity of a problem with such a threshold phenomenon is due to a unique type of instances. Therefore, understanding a threshold phenomenon is equivalent to understand intimately the difficulty of a problem. In the last few years, a renewed interest for this question gave rise to an impressive amount of work, as the bibliography testifies in the very case of SAT.

Let us now come to $k$-SAT formulae. The case $k = 2$ is now well known. As mentioned above, a threshold behavior (located at $\tau = 1$) has been proved [8,18]. Even the finest information has been recently achieved on the so-called scaling window [4]. Nevertheless the 2-SAT problem is known to belong to the class $\mathcal{P}$ and is therefore less symptomatic, which makes it slightly less interesting. On the contrary, any $k$-SAT problem with $k \geq 3$ is known to be $\mathcal{NP}$-complete and the threshold behavior would be of great interest. But nothing is proved in this direction, with the remarkable exception of Friedgut's paper [15]. And there are other positive facts. First, numerical experimentation shows that a threshold phenomenon is very likely [9,24,26]. Although not established, we write

$$\tau_0^{(k)}$$

for the prospective value of the transition point and a formula like $\tau_0^{(k)} \geq \alpha$ means only that for any $\tau < \alpha$, $k$-SAT formulae with $m = [\tau n]$ admit a solution with probability 1 (when $n$ tends towards $+\infty$). Second, some bounds on the probability for a random $k$-SAT formula to be satisfiable are known on some range of values for $\tau$. In [8], Chvatál and Reed prove that if $\tau < 2^{k-2}/k$, then almost every $k$-SAT formula is satisfiable: we shall therefore write

$$\tau_0^{(k)} \geq 2^{k-2}/k;$$

on the other side of the spectrum, the so-called first moment method yields an upper bound

$$\tau_0^{(k)} \leq -\log 2/\log(1-2^{-k})$$

above which almost every formula is unsatisfiable.

Concerning the special case of 3-SAT which attracts most of the interest as being the "easiest" unsolved problem of this type, more has been achieved. Chvatál and Reed's lower bound $2/3$ [8] has been considerably improved first to 1.63 [5] and then to 3.003 [16]. Much more recently the values 3.145 and 3.26 have been given [1,2]. As concerns upper bounds, the value 5.19 given by the first moment method has been improved to 5.08 in [12], to 4.76 in [22], then to 4.64 in [10], to 4.602 [23] and to 4.596 in [21]. The last result on this question comes from [11] where the value 4.506 is given. Numerical experiments tend to show that what is to be expected is something like

$$\tau_0^{(3)} \approx 4.25.$$

The threshold phenomenon, which is a basic feature in statistical and theoretical physics, leads naturally to a simple question. Is the $k$-SAT problem related in some sense to a physical phenomenon? In this way, researchers from other areas have given an efficient way to look at this problem [27]: a close connection between $k$-SAT and models from physics was established. See for example [25] where the method of replicas is presented. This fact is to be interpreted as an appeal to fruitful cross-cultural collaborations!

In this paper we introduce yet another drastically new viewpoint on the problem. It comes from additive number theory, a subarea of number theory concerned with solving equations over integers where the unknowns are linked by additive signs. The most representative problem of this kind is the famous Goldbach conjecture: every even integer can be written as the sum of two primes. Although quite astonishing at a first glance, this problem and the $k$-SAT problem are somehow related. Our approach uses the so-called circle method, the philosophy of which is the following: when you want to prove the existence of a solution to some given equation and if you have good reasons to believe that this number of solutions is "large", then just count these solutions. More precisely, we first have to get an integral formula for this number of solutions. Then, we need to separate the domain of integration into two parts (according to the original terms, the major and minor arcs): the first part (the major arcs) is intended to give the main contribution to the integral, the second part is intended to contribute only an erratum. Therefore, what remains to be done is to evaluate the integral on the major arcs or at least to bound it from below carefully, and to bound from above the contribution of the minor arcs. Naturally, on the one hand you have to try to make the major arcs as large as possible to catch an asymptotic information as precisely as possible (especially when you are looking for an asymptotic formula); on the other hand, the larger the major arcs are, the less precise the forthcoming computation will be: in conclusion, one of the difficulties of the method is to find a good balance.

In this paper, inspired by the circle method and its application by Freiman to some boolean problems [13,14], we propose an additive viewpoint on threshold properties of $k$-SAT.

In Section 2, we first show how to translate $k$-SAT into an additive problem (Proposition 1 and Corollary 2). In Section 3, using the basic principle of the circle method, we show how to evaluate the number of solutions to a $k$-SAT instance through an integral formula (Proposition 3). With these two results, we obtain an equivalent formulation of the threshold problem for $k$-SAT (Proposition 4). Section 4 is devoted to the special case $k = 3$, where some simplifications appear. Finally, our conclusion is a program to obtain conjectural (and hopefully proved) new lower bounds for $\tau_0^{(3)}$ (Section 5) and more generally for $\tau_0^{(k)}$.

We stress the fact that there is a huge difference between our approach and the preceding ones for obtaining lower bounds on $\tau_0^{(k)}$, namely we do not consider any particular algorithm

for which we try to show that it solves the problem in almost all cases (this approach naturally induces a restriction) but we try to prove directly that some kind of system has almost always a solution. No specific algorithm is used and thus, in principle, our method could lead to the true value (if it exists) of $\tau_0^{(k)}$.

## 2. Translating $k$-SAT into an additive problem

In this section, we explain our additive-number-theoretic point of view on the $k$-SAT problem. Proposition 1 and Corollary 2 below translate $k$-SAT into a linear boolean system of equations.

We consider an instance of $k$-SAT for an integer $k$ fixed. More precisely, we consider a system with $m$ clauses in $n$ variables, each clause depending exactly on $k$ variables.

The basic observation is the following proposition.

**Proposition 1.** *There is a one-to-one correspondence between the instances of $k$-SAT with $m$ clauses in $n$ variables and the set of $m$ boolean (that is $(x_1, \ldots, x_n) \in \{-1, 1\}^n$) linear forms of the form*

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \\ a_{m-1,1}x_1 + a_{m-1,2}x_2 + \cdots + a_{m-1,n}x_n \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n \end{cases} \tag{1}$$

*where*

- *for all $1 \le i \le m$, $1 \le j \le n$, $a_{i,j}$ belongs to $\{-1, 0, 1\}$,*
- *for all $1 \le i \le m$, the following relation holds*

$$\sum_{j=1}^{n} |a_{i,j}| = k. \tag{2}$$

*Moreover, through this correspondence, the satisfiability of a formula is equivalent to the existence of a solution to the system*

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \neq -k \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \neq -k \\ \vdots \quad\quad \vdots \quad\quad\quad\quad \vdots \quad\quad \vdots \\ a_{m-1,1}x_1 + a_{m-1,2}x_2 + \cdots + a_{m-1,n}x_n \neq -k \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n \neq -k. \end{cases} \tag{3}$$

**Proof.** The set of instances of $k$-SAT with $m$ clauses in $n$ variables (say $X_1, X_2, \ldots, X_n$) is exactly the set of formulae written as a conjunction of $m$ clauses (that is, in the Conjunctive Normal Form). Each of these $m$ clauses is a disjunction of literals (a variable or its negation), that is of the form ($1 \le i \le m$)

$$\tilde{X}_{\alpha_{i,1}} \vee \tilde{X}_{\alpha_{i,2}} \vee \cdots \vee \tilde{X}_{\alpha_{i,k}}, \tag{4}$$

where $\vee$ means "OR" and $\tilde{X}_{\alpha_{i,t}}$ is either $X_{\alpha_{i,t}}$ or its negation. The fact that there are exactly $k$ literals in (4) follows from the definition of $k$-SAT. Evidently, for any value of $i$, the $\alpha_{i,t}$'s ($1 \le t \le k$) are $k$ different integers in the range $\{1, 2, \ldots, n\}$.

Now, for a fixed $i$, the correspondence is the following: for any $j$ in $\{1, 2, \ldots, n\}$,

- if $j \in \{\alpha_{i,1}, \alpha_{i,2}, \ldots, \alpha_{i,k}\}$ then define $a_{i,j}$ as $+1$ if $\tilde{X}_j$ is the variable and $-1$ if it is its negation,
- otherwise, define $a_{i,j} = 0$.

So, by definition, $a_{i,j}$ belongs always to $\{-1, 0, 1\}$, and, since for all $1 \leq i \leq m$ there are exactly $k$ non-zero $\alpha_{i,t}$'s, (2) holds. The fact that this correspondence is one-to-one is immediate. This proves the first part of our assertion.

Now a formula is satisfiable if and only if there exists a choice of $X_1, X_2, \ldots, X_n$ which makes each of the $m$ clauses true. Consider, for any $1 \leq i \leq m$, the bijection which puts in correspondence $X_j$ and $x_j$ following the rule: if $X_j$ is "TRUE" then $x_j = 1$ and if $X_j$ is "FALSE" then $x_j = -1$.

Clearly, for any given value of $i$, $\tilde{X}_{\alpha_{i,1}} \vee \tilde{X}_{\alpha_{i,2}} \vee \cdots \vee \tilde{X}_{\alpha_{i,k}}$ is false if and only if $\tilde{X}_{\alpha_{i,t}}$ is false for any $t$ between 1 and $k$. Through the correspondence, this means that $a_{i,\alpha_{i,t}}$ is 1 if and only if $x_{\alpha_{i,t}}$ is $-1$ and vice versa; or equivalently that $a_{i,\alpha_{i,t}} x_{\alpha_{i,t}} = -1$ for any $t$. Since there are exactly $k$ non-zero $a_{i,j}$'s ($1 \leq j \leq n$) – namely the $a_{i,\alpha_{i,t}}$ for $t = 1, 2, \ldots, k$ – this is equivalent (for any $i$) to

$$a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = -k.$$

Hence the result. □

Since the values of each of the boolean linear form in (1) is an integer between $-k$ and $k$, we obtain the following corollary.

**Corollary 2.** *Through the above correspondence, satisfiability is equivalent to the existence of a solution to*

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = y_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = y_2 \\ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \qquad\quad \vdots \\ a_{m-1,1}x_1 + a_{m-1,2}x_2 + \cdots + a_{m-1,n}x_n = y_{m-1} \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n = y_m \end{cases} \tag{5}$$

*with*

$$(x_1, \ldots, x_n, y_1, \ldots, y_m) \in \{-1, 1\}^n \times \{-(k-1), -(k-2), \ldots, k-1, k\}^m. \tag{6}$$

Let us underline an important thing: in this corollary, we consider the solutions with

$$y_1, \ldots, y_m \in \{-(k-1), -(k-2), \ldots, k-1, k\}$$

but in fact, we can look at the $y_i$'s as elements of any subset $\mathcal{S}$ of $\mathbb{Z}$ which contains every value, but $-k$, that can be attained by the left-hand side of any equation of (1) (namely every integer less in absolute value than $k$ with the same parity as $k$). In other terms, we can take any $\mathcal{S}$ subject to

$$\{-(k-2), -(k-4), \ldots, k\} \subset \mathcal{S} \subset \mathbb{Z} \setminus \{-k\}. \tag{7}$$

We shall refer to such a set $\mathcal{S}$ in the following section.

Recall that our aim is to find a range (of the type $\tau < \tau_1^{(k)}$) such that for any value of $\tau$ in this range, as $n$ tends to infinity and $m = [\tau n]$, the number of such systems which have at least one solution is asymptotic to the total number of systems. This will lead to the lower bound

$$\tau_0^{(k)} \geq \tau_1^{(k)}.$$

We define $\mathcal{M}_{m,n}$ to be the set of matrices with $m$ lines and $n$ columns with coefficients $-1$, 0 or 1 with exactly $k$ non-zero coefficients on each line; we define $Y_m$ to be the set of $m$-dimensional vectors with integral coefficients in a set $\mathcal{S}$ satisfying (7) and $X_n$ to be the set of $n$-dimensional vectors with coefficients $-1$ or 1. The above-mentioned question can be restated as: being given a matrix $A$ in $\mathcal{M}_{m,n}$, solve

$$A\mathbf{X} = \mathbf{Y}, (\mathbf{X}, \mathbf{Y}) \in X_n \times Y_m. \tag{8}$$

## 3. The circle method and an integral formula

The circle method is a very powerful machinery to investigate number-theoretic problems with an additive flavor. It was introduced almost one century ago by Hardy and Ramanujan [19] in order to study the number of partitions of the integers. Since then, the method developed intensively [29] and succeeded in giving many beautiful results on hard problems like Goldbach's conjecture (every even number greater than 2 can be written as the sum of two primes) and Waring's problem (for any integer $k$, there is an integer $s$, depending only on $k$, such that every integer can be written as the sum of $s$ $k$-th powers). But these two famous and widely known problems are far from being the only cases in which the method was revealed fruitful and it would be hopeless to try to list all the possible applications. Anyway, Freiman began to apply the circle method in the so-called subset-sum problem, namely, given a set of integers $\mathcal{A} = \{a_1, \ldots, a_n\}$ and an integer $b$, determine whether or not $b$ can be written as the sum of the elements of some subset of $\mathcal{A}$. His approach then generated a lot of literature (among others [3,7,13,14]) to improve and generalize his work in higher dimensions ($a_i \in \mathbb{Z}^d$).

Here, the point is that the $k$-SAT problem, when translated into an additive language, is reminiscent of Freiman's viewpoint on the subset-sum problem. Nevertheless, there are dramatic differences, the main one being that, in the present case, the dimension ($m$ according to the notation of formula (1)) of the problem is not fixed! As shown in [28], in the subset-sum problem even a large (but fixed) dimension is a serious difficulty and the case is not completely solved (see the remarks on [6] in [28]).

We now come to the technical presentation of the tools required. As usual when the circle method is used, we begin by establishing an integral formula for the number of solutions to the system (5). Let us define

$$e(u) = \exp(2\pi \mathrm{i} u),$$

the basic fact is that for any integer $m$,

$$\int_0^1 e(mu)\mathrm{d}u = \delta_{m,0},$$

the Kronecker symbol. Consequently, $(x_1, \ldots, x_n)$ is a solution to the $i$-th equation of (5) – here the $a_{i,j}$'s and the $y_i$'s are considered as fixed – if and only if

$$\int_0^1 e\left((a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n - y_i)u\right)\mathrm{d}u = 1.$$

Thus the characteristic function of $(x_1, \ldots, x_n)$ being a solution to the system (5) is

$$\prod_{i=1}^{m} \int_0^1 e((a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n - y_i)u)\mathrm{d}u.$$

Hence, the total number of solutions $(x_1, \ldots, x_n, y_1, \ldots, y_m)$ which satisfy the instance determined by the matrix $A$ is

$$\begin{aligned} N(A) &= \sum_{\mathbf{X}\in X_n} \sum_{\mathbf{Y}\in Y_m} \prod_{i=1}^{m} \int_0^1 e\left((a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n - y_i)u_i\right) \mathrm{d}u_i \\ &= \int_{\mathbf{u}\in[0,1]^m} \sum_{\mathbf{X}\in X_n} \sum_{\mathbf{Y}\in Y_m} e(x_1\, \mathbf{A}_1 \cdot \mathbf{u}) \ldots e(x_n\, \mathbf{A}_n \cdot \mathbf{u}) e\left(-\sum_{i=1}^{m} y_i \cdot u_i\right) \mathrm{d}\mathbf{u} \\ &= \int_{\mathbf{u}\in[0,1]^m} \prod_{j=1}^{n} \left( \sum_{x_j\in\{-1,1\}} e(x_j\, \mathbf{A}_j \cdot \mathbf{u}) \right) \left( \sum_{\mathbf{Y}\in Y_m} e(-\mathbf{Y}\cdot\mathbf{u}) \right) \mathrm{d}\mathbf{u} \\ &= \int_{\mathbf{u}\in[0,1]^m} \prod_{j=1}^{n} \left( e(\mathbf{A}_j \cdot \mathbf{u}) + e(-\mathbf{A}_j \cdot \mathbf{u}) \right) \prod_{i=1}^{m} \left( \sum_{y_i\in\mathcal{S}} e(-y_i u_i) \right) \mathrm{d}\mathbf{u} \end{aligned}$$

where we have used the notation $\mathbf{A}_j$ to denote the $j$-th column of the matrix $A$, $\mathbf{u}$ for the vector with coordinates $(u_1, \ldots, u_m)$ and $\mathrm{d}\mathbf{u} = \mathrm{d}u_1 \ldots \mathrm{d}u_m$ (in the following, we need the notation $|\mathbf{u}|$ for the modulus of $\mathbf{u}$); the point simply denotes the canonical scalar product. We recall that $\mathcal{S}$ is any subset of $\mathbb{Z}$ satisfying (7) and may depend on $i$.

What we finally obtained is the following result which shows that there exists an exact integral formula giving the total number of solutions to a given $k$-SAT formula.

**Proposition 3.** *Let us consider an instance of $k$-SAT and define $A$ to be the matrix in $\mathcal{M}_{m,n}$ associated to this instance. Then, the number of assignments that make the instance true is equal to $N(A)$. In particular, the instance is solvable if and only if $N(A) > 0$.*

So, we may reformulate the threshold problem for $k$-SAT as follows.

**Proposition 4.** *Find (if it exists) a value $\tau_0^{(k)}$ such that*

- *for almost all $A \in \mathcal{M}_{m,n}$ with $m = [\tau n]$, $n \to +\infty$ and $\tau < \tau_0^{(k)}$, one has $N(A) > 0$,*
- *for almost all $A \in \mathcal{M}_{m,n}$ with $m = [\tau n]$, $n \to +\infty$ and $\tau > \tau_0^{(k)}$, one has $N(A) = 0$.*

More modestly, any value $\tau_1^{(k)}$ such that $\tau < \tau_1^{(k)}$ implies $N(A) > 0$ for almost all $A \in \mathcal{M}_{m,n}$ with $m = [\tau n]$, $n \to +\infty$, gives already the lower bound

$$\tau_0^{(k)} \geq \tau_1^{(k)},$$

for the prospective value of the transition point.

## 4. Specifying 3-SAT

For the sake of both simplicity of writing and clarity, we pursue further only in the special case $k = 3$ but there is nothing specific to this case. To go further in our computation, we first

need to choose a set $\mathcal{S}$. In the following, we shall take the "natural" choice

$$\mathcal{S} = \{-1, 1, 3\}. \tag{9}$$

We would like to stress the fact that there is absolutely no evidence for the optimality of this choice. We may have to change $\mathcal{S}$ later and we are free to do so as long as (7) is respected. Notice once again that we do not even need to assume that $\mathcal{S}$ is fixed. It could be needed in the future to consider $\mathcal{S}$ as a functional set depending on $m$ or even on each $i$ $(1 \leq i \leq m)$.

Anyway, with (9), $N(A)$ can be rewritten as

$$N(A) = \int_{\mathbf{u}\in[0,1]^m} \prod_{j=1}^{n} \left(e(\mathbf{A}_j.\mathbf{u}) + e(-\mathbf{A}_j.\mathbf{u})\right) \prod_{i=1}^{m} (e(u_i) + e(-u_i) + e(-3u_i))\, \mathrm{d}\mathbf{u}$$

and thus

$$\frac{N(A)}{2^n 3^m} = \int_{\mathbf{u}\in[0,1]^m} \prod_{j=1}^{n} \cos(2\pi \mathbf{A}_j.\mathbf{u}) \prod_{i=1}^{m} \left(\frac{1+2\cos(4\pi u_i)}{3}\right) e\left(-\sum_{i=1}^{m} u_i\right) \mathrm{d}\mathbf{u}. \tag{10}$$

Let us recall that our aim is to find a condition, of the form $\tau \leq \tau_1^{(3)}$ (with $m = [\tau n]$), implying that for almost all matrices $A \in \mathcal{M}_{m,n}$, $N(A) > 0$.

### 4.1. Some general considerations

As indicated above, the general strategy of the circle method is to separate the domain of integration in (10) into two parts depending on whether their contributions are large or small. Usually, in number theory, the major arcs (those parts of the domain where a large contribution is expected) are related to arithmetic properties: for example, in Waring's problem, the major arcs (there is a good deal of latitude in this choice) are small intervals centered around rational points with small denominators.

In the present situation, it is clear that the maximum of (the function to integrate)

$$\left| \prod_{j=1}^{n} \cos(2\pi \mathbf{A}_j.\mathbf{u}) \prod_{i=1}^{m} \left(\frac{1+2\cos(4\pi u_i)}{3}\right) e\left(-\sum_{i=1}^{m} u_i\right) \right|$$

are located in those points $\mathbf{u}$, all coordinates of which are integral and maybe some additional points with half-integral coordinates. The most direct idea is to work by defining the major arcs as the union of some neighborhoods of each of these points. We first regroup some of these points.

### 4.2. Symmetries in the integral

It is worth beginning by symmetrizing the integral. We let

$$\begin{aligned} I &= \frac{N(A)}{2^n 3^m} \\ &= \int_{\mathbf{u}\in[0,1]^m} \prod_{j=1}^{n} \cos(2\pi \mathbf{A}_j.\mathbf{u}) \prod_{i=1}^{m} \left(\frac{1+2\cos(4\pi u_i)}{3}\right) e\left(-\sum_{i=1}^{m} u_i\right) \mathrm{d}\mathbf{u}. \end{aligned}$$

The integral $I$ can be rewritten as

$$I = \int_{u_1\in[0,1/2]} \int_{(u_2,\ldots,u_m)\in[0,1]^{m-1}} + \int_{u_1\in[1/2,1]} \int_{(u_2,\ldots,u_m)\in[0,1]^{m-1}}.$$

Perform a change of variables in the second integral $t_1 = u_1 - 1/2$; we then obtain the same area of integration as in the first integral. Concerning the integrand, there appears a constant factor

$$-\prod_{j=1}^{n}(-1)^{a_{1,j}} = (-1)^{1+\sum_{j=1}^{n} a_{1,j}} = (-1)^{1+\sum_{j=1}^{n} |a_{1,j}|} = 1,$$

by (2). Repeating this process with $u_2, u_3$ up to $u_m$ and defining

$$J = \frac{I}{2^m},$$

we obtain

$$J = \int_{\mathbf{u}\in[0,1/2]^m} \prod_{j=1}^{n} \cos(2\pi \mathbf{A}_j \cdot \mathbf{u}) \prod_{i=1}^{m} \left(\frac{1+2\cos(4\pi u_i)}{3}\right) e\left(-\sum_{i=1}^{m} u_i\right) \mathrm{d}\mathbf{u}, \tag{11}$$

and now, for sure, any extremal value of the integrand takes place on a vertex of $[0, 1/2]^m$. At this point, it is good to consider the integral and to ask for more on the location of these extremal values. Clearly, a vertex of $[0, 1/2]^m$ corresponds to an extremum if and only if, for every $j$, $\mathbf{A}_j \cdot \mathbf{u}$ is integral. In any other case, $\mathbf{A}_j \cdot \mathbf{u}$ is half-integral and, in the vertex involved, the integrand is zero. What appears is that we shall need to examine the behavior of the integrand in the neighborhood of all the vertices of $[0, 1/2]^m$.

### *4.3. The job around the corner*

Introduce now some notations. For a vector $\mathbf{B} = (\beta_1, \ldots, \beta_m) \in \{0, 1\}^m$, define the cube

$$C_{\mathbf{B}} = \prod_{i=1}^{m} \left[\frac{\beta_i}{4}, \frac{1+\beta_i}{4}\right]$$

and the $\mathbf{B}$-conjugate, denoted by $\overline{\mathbf{x}}$, of a vector $\mathbf{x} = (x_1, \ldots, x_m)$ to be

$$\overline{x}_i = \begin{cases} x_i & \text{if } \beta_i = 0, \\ -x_i & \text{if } \beta_i = 1. \end{cases}$$

Separating the contributions of the $2^m$ different corners, we may rewrite $J$ as follows

$$J = \int_{\left[0,\frac{1}{2}\right]^m} = \sum_{\mathbf{B}\in\{0,1\}^m} \int_{C_{\mathbf{B}}} = \sum_{\mathbf{B}\in\{0,1\}^m} J_{\mathbf{B}}.$$

A new change of variables produces (the conjugate is in the sense of $\mathbf{B}$)

$$J_{\mathbf{B}} = \int_{\mathbf{u}\in\left[0,\frac{1}{4}\right]^m} \prod_{j=1}^{n} \cos(2\pi \overline{\mathbf{A}_j} \cdot \mathbf{u}) \prod_{i=1}^{m} \left(\frac{1+2\cos(4\pi u_i)}{3}\right) e\left(-\overline{\mathbf{1}} \cdot \mathbf{u}\right) \mathrm{d}\mathbf{u}. \tag{12}$$

Finally, we have

$$N(A) = 2^n 6^m \sum_{\mathbf{B}\in\{0,1\}^m} J_{\mathbf{B}},$$

and we are therefore led to the study of $2^m$ integrals of the same type.

## 5. A hope for the future supported by some observations and a research program

We cannot actually prove a new result on $\tau_0^{(3)}$. Instead, we would like to present some remarks and basic features which lead us to a research program. Once again, nothing below is specific to the case $k = 3$.

### 5.1. Some remarks

When we introduced the circle method, we wrote that the method works in those situations where it is believed that the number of solutions is "large". What about the present case? Let us imagine that the matrix $A$ is a fixed "generic" matrix. Any of the $m$ linear forms appearing in (1) takes the values $-3$ with probability $1/8$, $3$ with probability $1/8$, $-1$ with probability $3/8$ and $1$ with probability $3/8$. In an ideal case, these forms would be independent (this is of course not true) and the probability that none of them is $-3$ would be exactly $(7/8)^m$. Thus the expected number of solutions to (8) would be $2^n(7/8)^m$. We stress the fact that with this quantitative reasoning, writing that the number of expected solutions is 1 leads to $m/n = \log 2/\log(8/7)$, the value given by the first-moment approach, which is a positive sign .... Anyway, this rough reasoning leads to the idea that it could be reasonable to assume that the number of solutions $N(A)$ (for $A$ generic) is not too far from something of the form $\alpha^n \beta^m$ for some constants $\alpha$, $\beta$ and so that it may be true that, for certain matrices in $\mathcal{M}_{m,n}$, the number of corresponding solutions could be exponentially large in $n$. In particular, the circle method could apply. Our hope is that *most* matrices verify this assumption of genericity.

As seen above, for a given matrix $A$, we are led to compute the integrals $J_{\mathbf{B}}$. If we imagine that we take $A$ a "generic" matrix (in a sense to be made precise according to the forthcoming calculations), then for any $\mathbf{B}$ in $\{0,1\}^m$ (maybe not any, but again almost any — and we may adapt what follows), the vector $\overline{\mathbf{A}_j}$ should be generic.

Denote

$$\varphi(\mathbf{u}) = \prod_{j=1}^{n} \cos(2\pi \overline{\mathbf{A}_j} \cdot \mathbf{u}) \prod_{i=1}^{m} \left( \frac{1 + 2\cos(4\pi u_i)}{3} \right) e\left( -\overline{\mathbf{1}} \cdot \mathbf{u} \right),$$

the integrand in (12). On the domain of integration $[0, \frac{1}{4}]^m$, we have $|\varphi(\mathbf{u})| = 1$ if and only if $\mathbf{u} = \mathbf{0}$. On the more, if all the $u_i$ are non-zero, then $|\varphi(\mathbf{u})|$ should be exponentially small with $m$. Note that there are points at finite distance from $\mathbf{0}$ in which $\varphi$ is not small, for instance $(\frac{1}{4}, 0, \ldots, 0)$. Anyway, it is likely that the "large" values of $\varphi$ are located at some special places (major arcs).

### 5.2. A research program

Researching for these main contributions should be the heart of the problem.

To sum up, our hope is to prove that given a matrix $A \in \mathcal{M}_{m,n}$ which fulfills an additional property – a "generic" matrix – ($\star$), we may compute $J_{\mathbf{B}}$ (for a given $\mathbf{B}$) as follows: find a partition (depending on $A$, $\mathbf{B}$ and $m$, like everything below) of $[0, \frac{1}{4}]^m$ into three sets, say $\mathcal{V}_0$, $\mathcal{V}_1$ and $\mathcal{V}_2$ which are as follows:

- the set $\mathcal{V}_0$ will be the major arc (in particular it contains 0), so we have to be able to approximate precisely $\varphi$ on $\mathcal{V}_0$ and to compute asymptotically $\int_{\mathcal{V}_0} \varphi$ at the price of an erratum small enough,

- on the set $\mathcal{V}_1$ (the minor arc), the integrand $\varphi$ is so small that we can bound from above efficiently, so $\left|\int_{\mathcal{V}_1}\varphi\right| \leq \sup_{\mathcal{V}_1}|\varphi| = o\left(\int_{\mathcal{V}_0}\varphi\right)$,
- finally, the set $\mathcal{V}_2$ contains the remaining points; on $\mathcal{V}_2$, $\varphi$ may be large but (hopefully) $\mathcal{V}_2$ is small enough to make the method work using $\left|\int_{\mathcal{V}_2}\varphi\right| \leq \mathrm{mes}(\mathcal{V}_2) = o\left(\int_{\mathcal{V}_0}\varphi\right)$.

The conclusion would be that the complete integral behaves like its restriction on $\mathcal{V}_0$. This gives an estimate on $J_{\mathbf{B}}$. Then, collecting the contributions of all the $J_{\mathbf{B}}$'s, and hoping that not too much cancellations occur, we would get (something on) the behavior of $N(A)$ and prove $N(A) > 0$.

Finally, it remains only to show that the property $(\star)$ is fulfilled by almost all matrices $A$. This is tantamount to saying that the elements of $\mathcal{M}_{m,n}$ which fulfill $(\star)$ have density one as $m \to +\infty$.

The remarks below could help us to find what property $(\star)$, $\mathcal{V}_0$ and $\mathcal{V}_1$ could be.

*5.2.1. More on* $(\star)$

Typically, the "exceptional" matrices (those with zero density) for which the preceding calculations should not work could be a set of matrices with some exceptional "statistical properties" (too much 1's in the same column for instance).

*5.2.2. More on* $\mathcal{V}_1$

As an example, take

$$\mathcal{V}_1 = \left\{\mathbf{u} \in \left[0, \frac{1}{4}\right]^m : |\mathbf{u}| \geq c_0\sqrt{\log n}\right\}$$

and call $J_{\mathbf{B},1}$ the associated contribution in $J_{\mathbf{B}}$. We would like to underline the fact that the elements of $[0, 1/4]^m$ can have moduli up to $c\sqrt{m} \sim c\sqrt{\tau}\sqrt{n}$. Thus the restriction for $\mathcal{V}_1$ is far from trivial; on the contrary, most of the points of $[0, 1/4]^m$ do belong to $\mathcal{V}_1$.

Since $\mathcal{V}_1$ contains only points far from zero, we can majorize the modulus of $\varphi$ efficiently on this domain. Using some well-known inequalities, we get for some constant $c_1$:

$$\begin{aligned}|\varphi(\mathbf{u})| &\leq \exp\left(-c_1\sum_{i=1}^{m} u_i^2\right)\\ &\leq \exp(-c_1c_0^2\log n) = n^{-c_1c_0^2},\end{aligned}$$

from which we deduce

$$|J_{\mathbf{B},1}| \leq \frac{1}{4^m n^{c_1c_0^2}}.$$

*5.2.3. More on* $\mathcal{V}_0$

Concerning $\mathcal{V}_0$, we may proceed as in [28] and keep the same notation as on page 404 of this paper ($g$ is an at most quartic function). Denoting by $\|z\|$ the distance between a real $z$ and the set of integers $\mathbb{Z}$, one has

$$\begin{aligned}\cos(2\pi\overline{\mathbf{A}}_j\cdot\mathbf{u}) &= \cos(2\pi\|\overline{\mathbf{A}}_j\cdot\mathbf{u}\|)\\ &= \exp(-2\pi^2\|\overline{\mathbf{A}}_j\cdot\mathbf{u}\|^2)(1-g(\|\overline{\mathbf{A}}_j\cdot\mathbf{u}\|)).\end{aligned}$$

In the same way, one can approximate $(1 + 2\cos(4\pi u_i))/3$ with

$$\exp\left(-\frac{16}{3}\pi^2 u_i^2\right)$$

up to a multiplicative factor $1 + O(u_i^4)$.

Finally, the function $\phi$ can be approximate quite precisely as (here $i^2 = -1$)

$$\exp\left(-2\pi^2 \sum_{j=1}^{n} \|\overline{\mathbf{A}_j} \cdot \mathbf{u}\|^2 - \frac{16}{3}\pi^2 \sum_{k=1}^{m} u_k^2 + 2\pi \mathrm{i} \sum_{k=1}^{m} (1 - 2\beta_k) u_k\right).$$

Pushing the method up to the end should produce a main term (which would appear as the value of a certain Fourier transform of a quadratic form, as shown by the formula).

But a lot of work remains to be done in this direction.

## References


[1] D. Achlioptas, Setting 2 variables at a time yields a new lower bound for random 3-SAT (extended abstract), in: Proc. 32nd ACM Symp. on Theory of Computing, 2000, pp. 28–37.

[2] D. Achlioptas, G.B. Sorkin, Optimal myopic algorithms for random 3-SAT, in: Proc. 41st Symp. on Foundations of Computer Science, 2000, pp. 590–600.

[3] N. Alon, G. Freiman, On sums of subsets of a set of integers, Combinatorica 8 (1988) 297–306.

[4] B. Bollobás, C. Borgs, J.T. Chayes, J.H. Kim, D.B. Wilson, The scaling window of the 2-SAT transition, Random Structures Algorithms 18 (2001) 201–256.

[5] A.Z. Broder, A.M. Frieze, E. Upfal, On the satisfiability and maximum satisfiability of random 3-CNF formulas, in: Proc. of the 33rd IEEE Symposium in Foundations of Computer Science, SODA, 1993.

[6] M. Chaimovich, On solving dense *n*-dimensional subset sum problems, in: Proceedings of the Twenty-second Southeastern Conference on Combinatorics, Graph Theory, and Computing, Baton Rouge, LA, 1991, Congr. Numer. 84 (1991) 41–49.

[7] M. Chaimovich, G. Freiman, Z. Galil, Solving dense subset-sum problems by using analytical number theory, J. Complexity 5 (1989) 271–282.

[8] V. Chvátal, B. Reed, Mick gets some (the odds are on his side), in: Proc. of the Thirty-Third IEEE Symposium on Foundations of Computer Science, 1992, pp. 620–627.

[9] J. Crawford, L. Auton, Experimental results on the crossover point in random 3-SAT, Artif. Intell. 81 (1996) 31–57.

[10] O. Dubois, Y. Boufkhad, A general upper bound for the satisfiability threshold of random *r*-SAT formulae, J. Algorithms 24 (1997) 395–420.

[11] O. Dubois, Y. Boufkhad, J. Mandler, Typical random 3-SAT formulae and the satisfiability threshold (abstract), in: Proc. 11th ACM-SIAM Symp. on Discrete Algorithms 2000, pp. 126–127.

[12] A. El Maftouhi, W. Fernandez de la Vega, On random 3-sat, Combin. Probab. Comput. 4 (1995) 189–195.

[13] G.A. Freiman, New analytical results in subset-sum problem, in: Combinatorics and Algorithms (Jerusalem, 1988), Discrete Math. 114 (1993) 205–217.

[14] G. Freiman, On solvability of a system of two Boolean linear equations, in: Number Theory (New York, 1991–1995), Springer, New York, 1996, pp. 135–150.

[15] E. Friedgut, Sharp thresholds of graph properties, and the *k*-sat problem, with an appendix by Jean Bourgain, J. Amer. Math. Soc. 12 (1999) 1017–1054.

[16] A. Frieze, S. Suen, Analysis of two simple heuristics on a random instance of *k*-SAT, J. Algorithms 20 (1996) 312–355.

[17] M.R. Garey, D.S. Johnson, Computers and Intractability: A Guide to the Theory of $\mathcal{NP}$-Completeness, W. H. Freeman and Co, New-York, 1979.

[18] A. Goerdt, A threshold for unsatisfiability, in: Mathematical Foundations of Computer Science (Prague, 1992), in: Lecture Notes in Comput. Sci., vol. 629, Springer, 1992, pp. 264–274.

[19] G.H. Hardy, S. Ramanujan, Asymptotic formulae in combinatory analysis, Proc. London Math. Soc. 17 (1918) 75–115.

[20] B. Hayes, Can't get no satisfaction, Amer. Sci. 85 (1997) 108–112.

[21] S. Janson, Y.C. Stamatiou, M. Vamvakari, Bounding the unsatisfiability threshold of random 3-SAT, Random Structures Algorithms 17 (2000) 103–116.
[22] A. Kamath, R. Motwani, K. Palem, P. Spirakis, Tail bounds for occupancy and the satisfiability threshold conjecture, Random Structures Algorithms 7 (1995) 59–80.
[23] L.M. Kirousis, E. Kranakis, D. Krizanc, Y.C. Stamatiou, Approximating the unsatisfiability threshold of random formulas, Random Structures Algorithms 12 (1998) 253–269.
[24] T. Larrabee, Y. Tsuji, Evidence for a satisfiability threshold for random 3CNF formulas, in: Proc. of Spring Symposium on A. I. and $\mathcal{NP}$-hard Problems, Stanford, 1993, pp. 112–118.
[25] M. Mézard, Random systems and replica field theory, in: Géométries Fluctuantes en Mécanique Statistique et en Théorie des Champs (Les Houches, 1994), North-Holland, Amsterdam, 1996, pp. 1077–1090.
[26] D. Mitchell, B. Selman, H.J. Levesque, Hard and easy distributions of sat problems, in: Proc. of the Tenth National Conference on Artificial Intelligence, AAAI 1992, San Jose, pp. 459–465.
[27] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, L. Troyansky, $2+p$-SAT: relation of typical-case complexity to the nature of the phase transition, in: Statistical Physics Methods in Discrete Probability, Combinatorics, and Theoretical Computer Science, Princeton, NJ, 1997, Random Structures Algorithms 15 (1999) 414–435.
[28] A. Plagne, On the two-dimensional subset sum problem, Structure theory of set addition, Astérisque 258 (1999) 375–409.
[29] R.C. Vaughan, The Hardy-Littlewood Method, in: Cambridge Tracts in Mathematics, vol. 80, Cambridge University Press, Cambridge, New York, 1981, p. xi+172.